CSC 423 – A2

Devroop Banerjee
V00837868

1. **Exercise 2.2**
A monkey types on a 26-letter keyboard that has lowercase letters only. Each letter is chosen independently and uniformly at random from the alphabet. If the monkey types 1,000,000 letters, what is the expected number of times the sequence "proof" appears?

Exercise 2.2

$\frac{1}{26^5}$

$$x_i = \begin{cases} 1 & \text{if } i(i+1)\ldots(i+4) \rightarrow \text{has "proof"} \\ 0 & \text{otherwise} \end{cases}$$

$$x = \sum_{i=1}^{999996} x_i$$

$$E(x) = \sum E(x_i) = \underline{\underline{\frac{999996}{26^5}}}$$

2. **Exercise 2.18**
The following approach is often called reservoir sampling. Suppose we have a sequence of items passing by one at a time. We want to maintain a sample of one item with the property that it is uniformly distributed over all the items that we have seen at each step. Moreover, we want to accomplish this without knowing the total number of items in advance or storing all of the items that we see.

Consider the following algorithm, which stores just one item in memory at all times. When the first item appears, it is stored in the memory. When the kth item appears, it replaces the item in memory with probability 1/k. Explain why this algorithm solves the problem.

2> Exercise 2.18:

- Let $b_1, b_2, \ldots, b_n$ be the values of items observed at time b_t
- Let M_t be a r.v who's value is that of item in memory at time t.
- Need to show:

  (at any time t) $Pr[M_t = b_i] = 1/t \quad \forall 1 \leq i \leq t$.
- Base Case: (t = 1) $Pr[M_t = b_1] = 1$

  (At time t+1) $Pr[M_{t+1} = b_{t+1}] = 1/(t+1)$
- For $1 \leq i \leq t$, $Pr[M_{t+1} = a_i = Pr[$not replaced @ time t & $M_t = b_i]$

  $= Pr[$not replaced @ time t$] \cdot Pr[M_t = b_i]$

  $= \frac{t}{t+1} \times \frac{1}{t} = \boxed{\frac{1}{t+1}}$

3. Exercise 2.22

Let a1, a2, . . . , an be a list of n distinct numbers. We say that ai and aj are inverted if i < j but ai > aj . The Bubblesort sorting algorithm swaps pairwise adjacent inverted numbers in the list until there are no more inversions, so the list is in sorted order. Suppose that the input to Bubblesort is a random permutation, equally likely to be any of the n! permutations of n distinct numbers. Determine the expected number of inversions that need to be corrected by Bubblesort.

3> Exercise 2.22:

Let the r.v. X represent the number of inversions needed to sort the list. X is composed of a set of indicator r.v. X_{ij} where $\forall X_{ij} = 1$ when a_i & a_j are inverted. $(i < j ; a_i > a_j)$

We know: $X = \sum_{i<j} X_{ij} \dashrightarrow E[X] = \sum_{i<j} E[X_{ij}]$

(Since j is adjacent to i) $= \sum_{i=1}^{n} \sum_{j=i+1}^{n} Pr[X_{ij} = 1]$

Pr(Pair of numbers inverted) = 1/2 (a_i, a_j = random; equally likely of >)

Substituting back $Pr[X_{ij}] = 1/2$ ...

$E[X] = \sum_{i=1}^{n} \sum_{j=i+1}^{n} 1/2 - (\frac{n}{2})(\frac{n-1}{2})$

$\therefore$ Ans $\Rightarrow E[X] = \boxed{\frac{n(n-1)}{4}}$

4. Exercise 2.32

You need a new staff assistant, and you have n people to interview. You want to hire the best candidate for the position. When you interview a candidate, you can give them a score, with the highest score being the best and no ties being possible.

You interview the candidates one by one. Because of your company's hiring practices, after you interview the kth candidate, you either offer the candidate the job before the next interview or you forever lose the chance to hire that candidate. We suppose the candidates are interviewed in a random order, chosen uniformly at random from all n! possible orderings.

We consider the following strategy. First, interview m candidates but reject them all; these candidates give you an idea of how strong the field is. After the mth candidate, hire the first candidate you interview who is better than all of the previous candidates you have interviewed.

(a) Let E be the event that we hire the best assistant, and let E_i be the event that ith candidate is the best and we hire him. Determine $Pr(E_i)$, and show that

$$\mathbf{Pr(E)} = \frac{m}{n} \sum_{j=m+1}^{n} \frac{1}{j-1}$$

4) Exercise 2.32:

a) $Pr[E] = \sum_{i=1}^{n} Pr[E_i]$ * E_i are disjoint events *

For $i \leq m$, $Pr[E_i] = 0$ (none of the 1^{st} m candidates selected)

$i > m$, $Pr[E_i] = Pr[\text{best} = i^{th} \text{ candidate}] \cdot Pr[\text{chosen} = i^{th} \text{ candidate}]$

$= \frac{1}{n} \cdot Pr[\text{best of } (i-1) \text{ candidates} \in 1^{st} \text{ m candidates}]$

$= \frac{1}{n} \cdot \frac{m}{(i-1)}$

Therefore... $Pr[E] = \sum_{i=m+1}^{n} Pr[E_i] = \frac{m}{n} \sum_{i=m+1}^{n} \frac{1}{i-1}$ ■

(b) Bound $\sum_{j=m+1}^{n} \frac{1}{j-1}$ to obtain

$$\frac{m}{n} * (\ln n - \ln m) \leq Pr(E) \leq \frac{m}{n} * (\ln(n-1) - \ln(m-1))$$

b) Using Lemma:

- $Pr[E] \geq \frac{m}{n} \int_{m+1}^{n+1} \frac{1}{x-1} \, dx = \ln(x-1)\Big|_{m+1}^{n+1} = \frac{m}{n}(\ln(n) - \ln(m))$

&

- $Pr[E] \leq \frac{m}{n} \int_{m}^{n} \frac{1}{x-1} \, dx = \ln(x-1)\Big|_{m}^{n} = \frac{m}{n}(\ln(n-1) - \ln(m-1))$

(c) Show that m(ln n − lnm)/n is maximized when m = n/e, and explain why this means Pr(E) ≥ 1/e for this choice of m.

c) We need to find the "best" m. To do this, take the derivative (since bound from above is concave), set it equal to 0, and solve for m (this maximises Pr[E])

$$\frac{d}{dm}\frac{m}{n}(\ln(n)-\ln(m)) = \frac{\ln(n)}{n} - \frac{\ln(m)}{n} + \frac{1}{n} = 0$$

$$\ln(m) = \ln(n) - 1, \quad m = e^{\ln(n)-1} = e^{\ln(n)}e^{-1} = ne^{-1} = \frac{n}{e}$$

Substitute $m = n/e$ back into bound from (b).

$\longrightarrow \Pr[E] \geq \frac{1}{e}(\ln(n) - \ln(n/e)) = 1/e$

5. <u>Exercise 3.21</u>

(a) Chebyshev's inequality uses the variance of a random variable to bound its deviation from its expectation. We can also use higher moments. Suppose that we have a random variable X and an even integer k for which $E[(X - E[X])^k]$ is finite. Show that

$$\Pr\left(|X - E[X]| > t * \sqrt[k]{E[(X - E[X])^k]} \right) \leq \frac{1}{t^k}$$

5) Exercise 3.21:

a) Let $Y = (X - E[X])^k$

(By Markov's inequality) $P[Y > t^k E[Y]] \leq \frac{E[Y]}{t^k E[Y]} = \frac{1}{t^k}$

$\Rightarrow P[Y > t^k E[Y]] = P[\sqrt[k]{Y} > t \sqrt[k]{E[Y]}] = P[|X - E[X]| > t \sqrt[k]{E[(X-E[X])^k]}]$

The 1st step holds because we have taken k^{th} root of both sides of the inequality.

The 2nd step holds because when k is even, the k^{th} root of a number is the absolute value

Combining this with Markov's inequality gives:

$P[|X - E[X]| \geq t \sqrt[k]{E[(X-E[X])^k]}] \leq 1/t^k$

(b) Why is it difficult to derive a similar inequality when k is odd?

b) Since X is a r.v., $(X - E[X])^k$ might be negative for some k which has an odd value. Hence, Markov's inequality cannot be used here

6. **Exercise 3.22**

A fixed point of a permutation π [1, n] → [1, n] is a value for which $\pi(x) = x$. Find the variance in the number of fixed points of a permutation chosen uniformly at random from all permutations. (Hint: Let Xi be 1 if $\pi(i) = i$, so that $\sum_{i=1}^{n} X_i$ is the number of fixed points. You cannot use linearity to find $\mathrm{Var}[\sum_{i=1}^{n} X_i]$ but you can calculate it directly.)

6> Exercise 3.22:

Let x_i be an indicator r.v. for the event where $\pi(i) = i$ (this makes "i" a fixed point).

$x_i = \begin{cases} 1 & i \text{ is fixed point} \\ 0 & \text{otherwise.} \end{cases}$

Let $x = \sum_{i=1}^{n}$ be the number of fixed points.

$Var[x] = E[x^2] - (E[x])^2$ $\qquad E[x_i] = P[x_i] = 1/n$

→ $E[x] = E[\sum_{i=1}^{n} x_i] = \sum_{i=1}^{n} E[x_i] = \sum_{i=1}^{n} (1/n) = 1$

Now compute 1st term of variance.

$E[x^2] = E[(\sum_{i=1}^{n} x_i)^2]$

$= (\sum_{i=1}^{n} E[x_i^2]) + (\sum_{i=1}^{n} \sum_{j \neq i} E[x_i x_j])$

$= (\sum_{i=1}^{n} E[x_i]) + (\sum_{i=1}^{n} \sum_{j \neq i} E[x_i x_j])$ $\qquad (x_i^2 = x_i)$

$= 1 + (\sum_{i=1}^{n} \sum_{j \neq i} P[x_i = 1] E[x_i x_j \mid x_i = 1])$ $\qquad$ (Conditional exp. on $x_i = 1$)

$= 1 + (\sum_{i=1}^{n} \sum_{j \neq i} \frac{1}{n} \frac{1}{(n-1)})$ $\qquad$ * There are n−1 choices for mapping j, yielding 1/(n−1) as cond. prob. of fixed pt j * $\qquad (P[x_i = 1] = 1/n)$

$= 1 + 1 = 2$

Therefore, $Var[x] = 2 - 1 = 1$ (Ans)

7. **Exercise 3.26**

The weak law of large numbers states that, if $X_1, X_2, X_3, \ldots$ are independent and identically distributed random variables with mean μ and standard deviation σ, then for any constant $\varepsilon > 0$ we have

$$\lim_{n \to \infty} \Pr\left(\left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \mu \right| > \varepsilon \right) = 0$$

Use Chebyshev's inequality to prove the weak law of large numbers.

7> Exercise 3.26:

$Var\left(\frac{x_1 + x_2 + \ldots + x_n}{n}\right) = \frac{\sigma^2}{n}$

So, by Chebychev's inequality:

$P\left(\left|\frac{x_1 + x_2 + \ldots + x_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$

$\rightarrow 0$ as $n \rightarrow \infty$